



IFW16

RAW SEQUENCE LISTING

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:07

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

```

3 <110> APPLICANT: Smith, Daniel
5 <120> TITLE OF INVENTION: ALPHA-N-ACETYL GALACTOSAMINIDASE FROM CLOSTRIDIUM PERFRINGENS
7 <130> FILE REFERENCE: 0994.00134
9 <140> CURRENT APPLICATION NUMBER: 10/059,447B
10 <141> CURRENT FILING DATE: 2002-01-29
12 <150> PRIOR APPLICATION NUMBER: 60/064,683
13 <151> PRIOR FILING DATE: 1997-11-03
15 <150> PRIOR APPLICATION NUMBER: 10/185,476
16 <151> PRIOR FILING DATE: 1998-11-03
18 <160> NUMBER OF SEQ ID NOS: 12
20 <170> SOFTWARE: PatentIn version 3.2
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 29
24 <212> TYPE: PRT
25 <213> ORGANISM: Clostridium perfringens
28 <220> FEATURE:
--> 29 <221> NAME/KEY: Xaa
30 <222> LOCATION: (1)..(29)
31 <223> OTHER INFORMATION: Xaa is any protein
33 <220> FEATURE:
34 <221> NAME/KEY: misc_feature
35 <222> LOCATION: (18)..(18)
36 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
38 <220> FEATURE:
39 <221> NAME/KEY: misc_feature
40 <222> LOCATION: (29)..(29)
41 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
43 <400> SEQUENCE: 1
45 Met Lys Val Leu Gly Asn Tyr Ile Gln Arg Asn Phe His Tyr Asp Gly
46 1 5 10 15
--> 49 Lys Xaa Phe Tyr Thr Lys Gln Phe Asn Lys Pro Ile Xaa
50 20 25
53 <210> SEQ ID NO: 2
54 <211> LENGTH: 25
55 <212> TYPE: PRT
56 <213> ORGANISM: Clostridium perfringens
58 <400> SEQUENCE: 2
60 Lys Val Leu Gly Asn Tyr Ile Gln Arg Asn Phe His Tyr Asp Gly Lys
61 1 5 10 15
64 Ser Phe Tyr Thr Thr Ser Phe Leu Asn
65 20 25
68 <210> SEQ ID NO: 3
69 <211> LENGTH: 8

```



RAW SEQUENCE LISTING

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:07

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

70 <212> TYPE: PRT
71 <213> ORGANISM: Clostridium perfringens
73 <400> SEQUENCE: 3
75 Glu Asp Gly Ser Val Glu Val Lys
76 1 5
79 <210> SEQ ID NO: 4
80 <211> LENGTH: 7
81 <212> TYPE: PRT
82 <213> ORGANISM: Clostridium perfringens
84 <400> SEQUENCE: 4
86 Ala Thr Val Ser Leu Pro Arg
87 1 5
90 <210> SEQ ID NO: 5
91 <211> LENGTH: 8
92 <212> TYPE: PRT
93 <213> ORGANISM: Clostridium perfringens
95 <400> SEQUENCE: 5
97 Leu Pro Ala Ala Phe Arg Lys Ala
98 1 5
101 <210> SEQ ID NO: 6
102 <211> LENGTH: 7
103 <212> TYPE: PRT
104 <213> ORGANISM: Clostridium perfringens
106 <400> SEQUENCE: 6
108 Ile Ile Ile Leu Lys Glu Phe
109 1 5
112 <210> SEQ ID NO: 7
113 <211> LENGTH: 11
114 <212> TYPE: PRT
115 <213> ORGANISM: Clostridium perfringens
117 <400> SEQUENCE: 7
119 Asp Ser Gln Tyr Glu Phe Leu Ile Glu Arg
120 1 5 10
123 <210> SEQ ID NO: 8
124 <211> LENGTH: 9
125 <212> TYPE: PRT
126 <213> ORGANISM: Clostridium perfringens
128 <400> SEQUENCE: 8
130 Lys Tyr Asp Val Val Leu Gly Asn Lys
131 1 5
134 <210> SEQ ID NO: 9
135 <211> LENGTH: 6
136 <212> TYPE: PRT
137 <213> ORGANISM: Clostridium perfringens
139 <400> SEQUENCE: 9
141 Phe Pro Asn Glu Leu Lys
142 1 5
145 <210> SEQ ID NO: 10
146 <211> LENGTH: 19

RAW SEQUENCE LISTING

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:07

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

```

147 <212> TYPE: PRT
148 <213> ORGANISM: Clostridium perfringens
150 <400> SEQUENCE: 10
152 Ala Asn Phe Asn Gly Tyr Tyr Val Glu Leu Gly Gln Pro Ile Tyr Ala
153 1 5 10 15
156 Lys Ser Leu
160 <210> SEQ ID NO: 11
161 <211> LENGTH: 629
162 <212> TYPE: PRT
163 <213> ORGANISM: Clostridium perfringens
165 <400> SEQUENCE: 11
167 Met Lys Val Leu Gly Asn Tyr Ile Gln Arg Asn Phe His Tyr Asp Gly
168 1 5 10 15
171 Lys Ser Phe Tyr Thr Thr Ser Phe Leu Asn Pro Ile Leu Asn Glu Glu
172 20 25 30
175 Ile Leu Val His Thr Gln Asn Glu Phe Ile Ile Tyr Phe Val Asp Gly
176 35 40 45
179 Glu Ile Leu Pro Ser Ser Glu Arg Asn Asn Val Glu Ile Lys Lys Gln
180 50 55 60
183 Ser Glu Gln Leu Leu Val Val Asn Phe Ser Lys Asp Asn Leu Ser Val
184 65 70 75 80
187 Glu Val Asn Tyr Phe Val Glu Asn Lys Val Ile Asn Lys Lys Leu Thr
188 85 90 95
191 Val Phe Asn Cys Cys Lys Arg Ile Asn Tyr Ile Asp Cys Asp Thr Phe
192 100 105 110
195 Glu Phe Glu Asp Thr Asn Ile Tyr Tyr Pro Lys Lys Gln Asn Asn Ile
196 115 120 125
199 Glu Met Gly Asn Phe Asn Gly Tyr Tyr Val Leu Gly Gln Pro Ile Tyr
200 130 135 140
203 Ala Lys Ser Leu Phe Met Gly Met Glu Phe Pro Met Gly Glu Asn Arg
204 145 150 155 160
207 Ile Gln Glu Arg Lys Tyr Phe Ser Arg Tyr Tyr Gly Lys Ser Val
208 165 170 175
211 Glu Lys Arg Leu Asp Ile His Ser Ala Ile Ile Gly Ala Ala Pro Glu
212 180 185 190
215 Lys Ser Lys Glu Lys Ile Gln Ala Ser Phe Phe Glu Tyr Ile Lys Ala
216 195 200 205
219 Ile Ser Leu Pro Ala Thr Phe Arg Lys Gln Tyr Asn Ser Trp Tyr Asp
220 210 215 220
223 His Met Leu Asn Ile Thr Asn Asp Ser Ile Ile Lys Ser Phe Leu Glu
224 225 230 235 240
227 Ile Asn Arg Gly Phe Lys Asn Tyr Gly Ile Thr Leu Asp Ala Phe Val
228 245 250 255
231 Val Asp Asp Gly Trp Ala Asn Tyr Glu Ser Val Trp Glu Phe Asn Asp
232 260 265 270
235 Lys Phe Pro Asn Glu Leu Lys Asp Ile Ser Glu Cys Val Lys Asn Leu
236 275 280 285
239 Gly Ser Thr Leu Gly Leu Trp Ile Gly Pro Arg Gly Gly Tyr Asn Gly
240 290 295 300

```

RAW SEQUENCE LISTING

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:07

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

```

243 Thr Gln Val Thr Met Ser Asp Trp Leu Glu Lys Asn Lys Asp Leu Asn
244 305 310 315 320
247 Ile Gly Ser Lys Asn Lys Ile Ser Asn Asp Val Asn Val Gly Asp Phe
248 325 330 335
251 Asn Tyr Leu Arg Lys Arg Asn Lys Glu Lys Met Leu Glu Tyr Gln Ser
252 340 345 350
255 Lys Tyr Asp Ile Ser Tyr Trp Lys Ile Asp Gly Met Leu Leu Lys Pro
256 355 360 365
259 Asp Thr Glu Asp Glu Ser Gly Pro Tyr Gly Met His Thr Met Thr Ala
260 370 375 380
263 Val Tyr Glu Phe Met Ile Ser Leu Phe Asn Glu Leu Arg Glu Glu Arg
264 385 390 395 400
267 Gly Glu Lys Ser Phe Trp Ile Asn Leu Thr Ser Tyr Val Asn Pro Ser
268 405 410 415
271 Pro Trp Phe Leu Lys Trp Val Asn Ser Leu Trp Ile Gln Thr Ser Gln
272 420 425 430
275 Asp Val Gly Phe Thr Pro Asn Gly Glu Asn Asp Ile Gln Lys Met Ile
276 435 440 445
279 Thr Tyr Arg Asp Ser Gln Tyr Trp Glu Phe Leu Ile Glu Arg Asp Ile
280 450 455 460
283 Gln Leu Pro Leu Cys Ser Leu Tyr Asn His Glu Pro Ile Tyr Ala Glu
284 465 470 475 480
287 Ser Ala Ser Met Trp Tyr Leu Asp His Gln Ile Tyr Cys Ser Ile Glu
288 485 490 495
291 Glu Ile Phe Lys Glu Tyr Leu Met Phe Ile Ala Thr Arg Gly Asn Ala
292 500 505 510
295 Phe Trp Glu Phe Tyr Tyr Ser Tyr Ser Met Phe Asp Asp Glu Arg Trp
296 515 520 525
299 Glu Val Asn Ala Gln Ala Ile Lys Trp Ile Glu Glu Asn Tyr Pro Ile
300 530 535 540
303 Leu Lys Asn Ser Thr Phe Gly Thr Lys Pro Ser Leu Met Gly Val
304 545 550 555 560
307 Tyr Gly Tyr Tyr Cys Gln Ser Asp Ser Gly Ser Lys Ser Ile Ile Ser
308 565 570 575
311 Phe Arg Asn Pro Ser Asp Glu Ile Lys Ser Tyr Lys Leu Glu Asn Ile
312 580 585 590
315 Glu Pro Lys Lys Tyr Asp Val Val Leu Gly Asn Lys Asn Tyr Lys Val
316 595 600 605
319 Phe Glu Asp Gly Ser Val Glu Val Lys Leu Asn Pro Lys Glu Ile Ile
320 610 615 620
323 Ile Leu Lys Ser Lys
324 625
327 <210> SEQ ID NO: 12
328 <211> LENGTH: 1890
329 <212> TYPE: DNA
330 <213> ORGANISM: Clostridium perfringens
332 <400> SEQUENCE: 12
333 atgaaagtat taggaaatta tattcaaaga aatttttcatt atgatggaaa aagtttttat 60
335 accacatcat ttttaaatec tattctaaat gaagaaatat tagttcatat acaaaatgaa 120

```

RAW SEQUENCE LISTING

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:07

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

337	tttattatct	attttgtaga	tggagaaata	ttaccttctt	ctgagatgaa	tgtggagatt	180
339	aagaagcaaa	gtgaacaact	tttagtggtg	aattttagta	aagataactt	atctgttgaa	240
341	gttaattatt	ttgtggaaaa	taagggtata	aataaaaagc	taacagtttt	caattgttgt	300
343	aaacgtatta	atttatattga	ctgtgatact	tttgaatttg	aggatactaa	taatatctat	360
345	taccctaaaa	aacagaataa	tataaaggaa	atggggaatt	ttaacggata	ctatgtagaa	420
347	ttagggcaac	ctatttatgc	aaaatcttta	ttcatgggaa	tgggaatttc	tatgggagaa	480
349	aatcgtattc	aagaaagaaa	gtatttttca	aggtattatt	atggaaaaag	tgtagaaaaa	540
351	agattagata	tacattcagc	aattattgga	gctgctccag	aaaaatcaaa	agaaaaaatt	600
353	caagcttcat	tttttgagta	tattaaagct	atatctttgc	cagctacttt	tagaaaacag	660
355	tataattctt	ggtatgatca	tatgctaaac	attactaatg	atagcataat	aaaaagtttc	720
357	ttagaaataa	atagaggctt	taaaaactat	ggaattactt	tagatgcctt	tgtagttgat	780
359	gatggttggg	ctaattatga	aagtgtttgg	gaatttaatg	ataagtttcc	taatgaatta	840
361	aaagatatat	cagaatgtgt	aaaaaatctt	ggttcaactt	taggactatg	gattggtcca	900
363	cgtggtggat	ataatggaac	tcaagttact	atgagtgatt	ggtagaaaaa	aaataaggat	960
365	ttaaacatag	gatctaaaaa	taaaatttct	aatgatgtaa	atgtaggaga	ctttaattat	1020
367	cttagaaaaga	tgaagaaaaa	aatgttagag	taccaaagca	aatatgacat	ctcctattgg	1080
369	aaaattgatg	gaatgttatt	aaagccagat	actgaggatg	aaagtggacc	atatggtatg	1140
371	catactatga	cggcagtata	tgaatttatg	attagtctat	ttaatgagtt	aagagaagaa	1200
373	agaggagaaa	agagtttttg	gatcaatctt	acatcttatg	ttaatcctag	cccttggttt	1260
375	ttaaagtggg	taaatagtct	ttggattcag	acttcacaag	atgttggcct	tactccaaat	1320
377	ggaggaaatg	atattcagaa	aatgatcaca	tatcgtgatt	ctcaatatta	tgaattcttg	1380
379	attgaaagag	atattcaact	tccattatgt	agcttatata	atcatgaacc	tatttatgca	1440
381	gagtctgcaa	gtatgtggta	tttagatcat	caaatctatt	gttctataga	agagttaa	1500
383	gagtatttaa	tgtttattgc	tactcgtgga	aatgcttttt	gggaatttta	ttattcttat	1560
385	tccatgtttg	atgatgaacg	ttgggaagta	aacgcacaag	ccattaagtg	gattgaggaa	1620
387	aattatccaa	tattaaaaaa	tagtactttc	tttggaaaca	agcctagcct	tatgggagta	1680
389	tatggatact	attgtcaatc	agattctggg	tcaaaatcaa	ttatttcatt	tagaaaccca	1740
391	tcagatgaaa	ttaaacttta	taaacttgag	aatatagaac	caaagaaata	tgacgtagtt	1800
393	ctaggcaata	aaaattataa	agtttttgaa	gatggttccg	ttgaagttaa	attaaatcct	1860
395	aaagaaatta	ttatacttaa	gagtaaataa				1890

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/059,447B

DATE: 07/23/2004
TIME: 12:46:08

Input Set : A:\DAN SMITH.ST25.txt
Output Set: N:\CRF4\07232004\J059447B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 18,29

VERIFICATION SUMMARY

DATE: 07/23/2004

PATENT APPLICATION: US/10/059,447B

TIME: 12:46:08

Input Set : A:\DAN SMITH.ST25.txt

Output Set: N:\CRF4\07232004\J059447B.raw

L:29 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1

L:49 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16